

Claims

- 1 1. A method of data analysis comprising,
2 employing data comprising, a plurality of records, each of said records having an
3 associated plurality of attributes, said plurality of records being divisible into at least two
4 categories,
5 providing a multidimensional array having intersecting first and second axes,
6 assigning each of said attributes as a one dimensional vector aligned along said first axis,
7 assigning each of said records as a one dimensional vector aligned along said second axis,
8 displaying a graphical indication of at least one of an occurrence and a value of each said
9 attribute for each said record at an intersection of each said record vector with each said attribute
10 vector, and
11 manipulating at least one of said record vectors and said attribute vectors to produce a
12 graphical pattern representative of said at least two categories.
- 1 2. The method of claim 1, wherein said manipulating step comprises transposing vectors in
2 at least one of a pair of said record vectors and a pair of said attribute vectors to produce said
3 graphical pattern representative of said at least two categories.
- 1 3. The method of claim 1 further comprising, determining, from said graphical pattern, a
2 result-effective subset of attributes that is sufficient to divide said records into said at least two
3 categories.
- 1 4. The method of claim 3, wherein said result-effective subset is a minimum subset.
- 1 5. The method of claim 4, wherein said minimum subset is unique.
- 1 6. The method of claim 2 further comprising, repeating said transposing step until said
2 graphical pattern is produced.

- 1 7. The method of claim 1, wherein said graphical indication has at least two states.
- 1 8. The method of claim 7, wherein each of said states is represented by a color.
- 1 9. The method of claim 7, wherein each of said states is represented by a symbol.
- 1 10. The method of claim 7, wherein each of said states is represented by an integer.
- 1 11. The method of claim 7, wherein said each of said states is represented by a shade of gray
- 1 12. The method of claim 1, wherein said data comprises training data obtained from control specimens.
- 1 13. The method of claim 1, wherein said manipulating step comprises employing a principal uncorrelated record set algorithm.
- 1 14. The method of claim 1, wherein said records represent cells and said attributes are properties of said cells.
- 1 15. The method of claim 1, wherein said records represent mammals and said attributes are characteristics of said mammals.
- 1 16. The method of claim 1, wherein said records represent a sample from a mammal and said attributes are biological markers.
- 1 17. The method of claim 16, wherein said biological marker is a gene product.
- 1 18. The method of claim 16, wherein said biological marker is at least one of a protein and an mRNA.

1 19. The method of claim 1, wherein at least one of said at least two categories represents a
2 predisposition to contract a disease.

1 20. The method of claim 19, wherein said disease is leukemia.

1 21. The method of claim 1, wherein at least one of said at least two categories represents a
2 predisposition to a medical treatment efficacy.

1 22. The method of claim 1, wherein a first category represents a mammal having a first
2 phenotype and a second category represents a mammal having a second, different phenotype.

1 23. The method of claim 22, wherein the first phenotype is a disease affected phenotype.

1 24. The method of claim 22, wherein the second phenotype is a non-disease affected
2 phenotype.

1 25. The method of claim 22, wherein the disease is a cancer.

1 26. The method of claim 12, wherein said control specimens are mammals having a non-
2 disease affected phenotype.

1 27. The method of claim 12, wherein said control specimens are mammals having a disease
2 affected phenotype.

1 28. A system for data analysis, comprising;
2 a processor adapted for,
3 providing data including, a plurality of records, each of said records having an
4 associated plurality of attributes, said plurality of records being divisible into at least two
5 categories,
6 providing a two dimensional array having intersecting first and second axes,

7 assigning each of said attributes as a one dimensional vector aligned along said
8 first axis,

9 assigning each of said records as a one dimensional vector aligned along said
10 second axis, and

11 manipulating at least one of said record vectors and said attribute vectors to
12 produce a graphical pattern representative of said at least two categories; and

13 a display adapted to display a graphical indication of at least one of an occurrence and a
14 value of each said attribute for each said record at an intersection of each said record vector with
15 each said attribute vector.

29. The system of claim 28, wherein said manipulating comprises transposing vectors in at
least one of a pair of said record vectors and a pair of said attribute vectors to produce said
graphical pattern representative of said at least two categories.

30. The system of claim 28, wherein said processor is further adapted for determining, from
said graphical pattern, a result-effective subset of attributes that is sufficient to divide said
records into said at least two categories.

31. The system of claim 30, wherein said result-effective subset is a minimum subset.

1 32. The system of claim 30, wherein said minimum subset is unique.

1 33. The system of claim 29, wherein said processor is further adapted for repeating said
2 transposing step until said graphical pattern is produced.

1 34. The system of claim 28, wherein said graphical indication has at least two states.

1 35. The system of claim 34, wherein each of said states is represented by a color.

1 36. The system of claim 34, wherein each of said states is represented by a symbol.

- 1 37. The system of claim 34, wherein each of said states is represented by an integer.
- 1 38. The system of claim 34, wherein said each of said states is represented by a shade of gray.
- 1 39. The system of claim 28, wherein said data comprises training data obtained from control
2 specimens.
- 1 40. The system of claim 28, wherein said manipulating comprises employing a principal
2 uncorrelated record set algorithm.
- 1 41. The system of claim 28, wherein said records represent cells and said attributes are
2 properties of said cells.
- 1 42. The system of claim 28, wherein said records represent mammals and said attributes are
2 characteristics of said mammals.
- 1 43. The system of claim 28, wherein said records represent a sample from a mammal and said
2 attributes are biological markers.
- 1 44. The system of claim 43, wherein said biological marker is a gene product.
- 1 45. The system of claim 43, wherein said biological marker is at least one of a protein and an
2 mRNA.
- 1 46. The system of claim 28, wherein at least one of said at least two categories represents a
2 predisposition to contract a disease.
- 1 47. The system of claim 46, wherein said disease is leukemia.

1 48. The system of claim 28, wherein at least one of said at least two categories represents a
2 predisposition to a medical treatment efficacy.

1 49. The system of claim 28, wherein a first category represents a mammal having a first
2 phenotype and a second category represents a mammal having a second, different phenotype.

1 50. The system of claim 49, wherein the first phenotype is a disease affected phenotype.

1 51. The system of claim 49, wherein the second phenotype is a non-disease affected
2 phenotype.

52. The system of claim 50, wherein the disease is a cancer.

53. The system of claim 39, wherein said control specimens are mammals having a non-
disease affected phenotype.

54. The system of claim 39, wherein said control specimens are mammals having a disease
affected phenotype.

55. A computer program recorded on a computer-readable medium for graphical data
analysis, said computer program when operating performing said steps of,
accepting data comprising, a plurality of records, each of said records having an
associated plurality of attributes, said plurality of records being divisible into at least two
categories,

constructing a multidimensional array having intersecting first and second axes,
assigning each of said attributes as a one dimensional vector aligned along said first axis,
assigning each of said records as a one dimensional vector aligned along said second axis,
displaying a graphical indication of at least one of an occurrence and a value of said
attribute for each said record at an intersection of each said record vector with each said attribute
vector, and

12 manipulating at least one of said record vectors and said attribute vectors to produce a
13 graphical pattern representative of said at least two categories.

1 56. The computer program of claim 55, wherein said manipulating step comprises
2 transposing vectors in at least one of a pair of said record vectors and a pair of said attribute
3 vectors to produce said graphical pattern representative of said at least two categories.

1 57. The computer program of claim 55, when operating, further comprising, determining,
2 form said graphical pattern, a result-effective subset of attributes that is sufficient to divide said
3 records into said at least two categories.